AUTOMATED INTERICTAL EPILEPSY DETECTION AND SOURCE LOCALIZATION IN EEG

S. Raiesdana[1], A. M. Nasrabadi[2]
[1] Department of electrical and computer engineering, Islamic Azad University, Qazvin Branch, Qazvin, Iran
[2]Department of electrical and computer engineering, Shahed University, Tehran, Iran
e-mail: raies_dana@yahoo.com

Abstract-**Multichannel recording of electroencephalography from neuronal currents in the brain generates large amount of data about the functionality of the brain. Epilepsy has distinct characteristic waveforms in the EEG signal during seizure and between them. But since manually analyzing large amount of data produces fatigue and error, automated technique for epilepsy detection and localization is an essence. Independent component analysis (ICA) is a technique which extracts statistical independent components from a set of measured signal and the Constrained ICA is a new framework to incorporate prior information in the form of constraints into the ICA contrast function.**
In this paper the cICA technique with an algorithm for source localization, Recursively Applied and Projected Multiple Signal Classification (RAP-MUSIC), were applied to an epileptic EEG data. The results show a superior separation of epileptic components with lower error (measured as performance index) and higher PSNR compared to conventional methods.
***Keywords* - Constrained Independent Component Analysis, epileptiform discharges, spike, spike - wave, Source localization**

I. INTRODUCTION

Epilepsy, the second most popular brain disease after stroke, is the appearance of a sudden and transient perturbation in the brain function and a result of the synchronous discharges of brain neurons. During seizure, the scalp EEG of patients with epilepsy is characterized by high amplitude and synchronized periodic EEG waveforms. Also in between seizures, the epileptiform transient waveforms which include spikes and sharp waves, are typically observed in the scalp EEG of such patients. Hence investigation of the EEG signal can involve valuable information of such disorders.

Isolating and classification of such waveforms with visual scanning of EEG is a hard and time consuming labor which needs an expert, specially in continuous and long term monitoring of the EEG, no online detection scheme exists. In addition, because epileptic discharges occur occasionally, they may be ignored in the reviewing process. Hence, a need arises for a reliable and automated detection scheme for detecting the onset of epileptic discharges [7].

Unfortunately automatic spike detection is difficult for number of reasons. In fact two human experts often do not mark the same events as spikes (or sharp waves) and the ratio of candidate spike events to actual spike events (false positives) is very large due to the existence of numerous noises and artifacts from other biosignals as EMG, EOG and ECG. Furthermore spike morphology and background vary widely between patients and well defined training sets are time consuming and expensive to develop.

But since EEG is recorded by a particular spatial distribution of electrodes on the scalp or within the brain, rather than time, it depends on space. This characteristic pushes the EEGer looks at all EEG channels and explores for particular patterns with specific spatial distribution. Mimicking this method (using the spatial and temporal information of all channels) in automatic detecting systems can eliminate the false positives. In this paper a new spatiotemporal analysis, constrained ICA, is used for epileptiform discharges detection.

However, in acute cases of disease which need surgery, the presurgical assessment of epileptic patients, is a multi stage process in which the ultimate aim is to determine the focal epileptic regions. This focal region is often taken out in the surgery. Hence in spite of detecting epileptic sources, assigning their location within the brain is an essence too. So the second part of our automatic detecting algorithm is source localization. EEG source localization (finding the location of equivalent current dipoles in the human brain from measurement of scalp potentials) is a suitable method for mapping the extent of abnormal cortex and to focus intracranial electrodes.

This localization problem is commonly referred to as the inverse problem of electroencephalography. Although the inverse problem is ill posed as there is an infinite number of source configurations inside the brain that can produce the exactly the same potentials at the surface of the head. To overcome this difficulty it must be assumed that the sources are made of current dipoles with a small number of dipoles and a sufficient number of electrodes. In this sense dipolar model is a reasonable approximation for focal sources but still different combination of sources can produce very similar patterns [6]. A widely used approach for source localization is based on PCA. For instance, the nonzero singular values used in MUSIC[1] algorithm for defining a signal subspace, has some advantages over direct least square methods in which all sources are located simultaneously. Also searching over the parameter space avoids local minima encountered in searching for multiple sources over a nonconvex error surface.

Here the source localization problem is dealt with a new version of MUSIC algorithm, recursively applied and projected MUSIC [6], in which each source is found as the global maximizer of a different cost function.

II. METHODOLOGY

A. Constrained ICA (cICA)

Epileptiform discharges are mixed with the background activity and sometimes with artifacts as EMG- EOG- ECG. Independent component analysis (ICA) is a new technique for separating statistically independent components from a mixture of data by high order statistics. ICA does not require any assumption regarding generator model in the head, but some underlying assumptions related to statistical

[1] Multiple Signal Classification

characteristic of source distribution are needed for implementing ICA. These assumptions are such that scalp EEG is a linear summation of the electrical activity from various brain regions and also potential field distribution is spatially fixed and only the electrical strength is changing within these regions. Finally it is assumed that any epileptic activity is independent of any ongoing background EEG [4].

Several different implementations of ICA can be found in literature. Most of them estimate same number of ICs as the channels of observed mixture data. Though often in real application, the number of components needed to be recovered is less. Especially in EEG data, the number of components of interest is much fewer than the observed data consist of a large number of components related to noise and artifacts.

In these cases some pre or post processing algorithms must be integrated to algorithm in order to reduce the dimensionality. Preprocessing may distort data and destroy week components and also in post processing additional time and resources are needed. In contrast, it seems that constrainting the ICA [3] by incorporating some assumptions and priory information such as statistical properties or rough templates of the sources is helpful. The cICA algorithm seeks those components statistically independent from others and closest to some reference signal in some closeness measure.

Let us denote the time varying observed signal by $\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_n(t)]^{\mathrm{T}}$ and the sources signals consisting of independent components (ICs) by $\mathbf{s}(t) = [s_1(t)\ s_2(t)\ \ldots s_m(t)]^{\mathrm{T}}$, so

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{1}$$

where the matrix A is a mixing matrix of size $n \times m$. In ICA, the aim is to determine the demixing matrix $\mathbf{W}$, such that $\mathbf{s}(t) = \mathbf{W}\mathbf{x}(t)$.

The optimization problem for extracting ICs is a contrast function in which the negentropy criteria (the below definition) is maximized.

$$J(\mathbf{y}) = H(\mathbf{y}_{gaus}) - H(\mathbf{y}) \tag{2}$$

$H(.)$ is differential entropy and $\mathbf{y}_{gaus}$ is a Gausian random variable. Maximizing the negentropy, as to central limit theory, produces a single IC.

In general due to the above definition, the cICA deals with the following constrained minimization problem [2]:

$$\begin{aligned} &\min imize \quad C(\mathbf{y}) \\ &subject \quad to \quad \mathbf{g}(\mathbf{y}:W)\ and\ /\ or\ \mathbf{h}(\mathbf{y}:W) \end{aligned} \tag{3}$$

where C(.) represents contrast function, and g and h define the vectors of inequality and equality constraints respectively.

B. RAP-MUSIC

As mentioned previously in order to improve the specificity of the detection procedure, the subset of the detected spikes captured in cICA, can be localized as focal neuronal sources. For doing this an equivalent current dipole model is fitted to the data in vicinity of spikes and the solution is accepted only if there is at least a %95 fit of a dipole model to the data. RAP-MUSIC algorithm is used to localize the equivalent current dipoles from the spatiotemporal data by using all of the spatial information instead of individual scalp maps. As seen in the last section, ICA can produce temporal components besides their corresponding spatial topographies. So we can form a lower rank approximation of $x(t)$ as $y(t) = \sum_m w_m s_m$.

where $m \in \mathrm{N}$ and N is a set of spiky components obtained by cICA.

In RAP-MUSIC algorithm, a signal subspace is including all of the spatial information of epileptic patterns. The space with residual information forms the noise subspace [6].

The signal subspace is found from SVD[2] of y as:

$$\mathbf{R}_y = \mathbf{\Phi\Lambda\Phi}^{\mathrm{T}} = [\mathbf{\Phi}_s\ \mathbf{\Phi}_n]\begin{bmatrix}\mathbf{\Lambda}_s & 0\\ 0 & \mathbf{\Lambda}_n\end{bmatrix}[\mathbf{\Phi}_s\ \mathbf{\Phi}_n]^{\mathrm{T}} = \mathbf{\Phi}_s\mathbf{\Lambda}_s\mathbf{\Phi}_s^{\mathrm{T}} + \mathbf{\Phi}_n\mathbf{\Lambda}_n\mathbf{\Phi}_n^{\mathrm{T}} \tag{4}$$

where the columns of $\Phi_s\ and\ \Phi_n$ spans respectively the estimated signal and noise subspaces. RAP-MUSIC finds sources in an automated and recursive form by projecting the topography for candidate source locations against the estimated signal subspace.

III. Simulation

A. Data preparation

In order to evaluate the capacities of the two algorithms explained above, an epileptic EEG signal is simulated. In this sense the epileptifom discharges generated by computer were superimposed on the background activity from a normal subject. This model includes a dipole with fixed location and orientation and a moment having a spike- like waveform [5]. The epileptic waveform is computed at each electrode location using the potential field obtained by solving the forward problem for the dipole model with a spherical 3 layer head model. Each artificially made epileptiform discharge is a summation of two transients represented by the moments of the dipoles. Epileptiform activity1 (EA1) is a 'spike-wave' and the epileptiform activity2 (EA2) is a 'spike' with known time varying moments. Also for assessing the ability of specificity, two dipoles are put close to each other (fig. 1) and then the potential distributions of them are calculated at defined electrode locations.

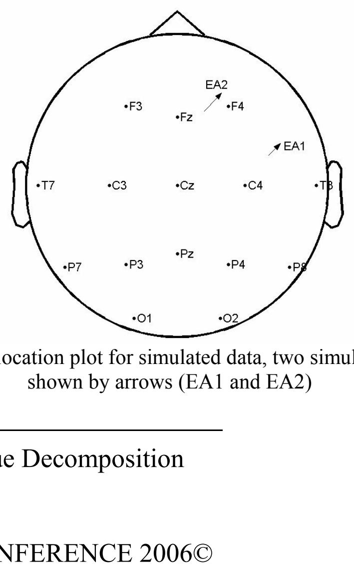

Fig. 1. Electrode location plot for simulated data, two simulated dipoles are shown by arrows (EA1 and EA2)

[2] Singular Value Decomposition

B. Algorithm adjustments

We use $g(\mathbf{w}) = \varepsilon(y, r) - \xi$ as the inequality constraint for cICA, where $\varepsilon(y, r)$ is the closeness measure between the output $\mathbf{y}$ and the reference signal $\mathbf{r}$ and ξ is some closeness threshold. In this paper the measure of closeness is captured by correlation and ξ is set to one. The relative morphology of the reference signal and the time course of it can be achieved simply and automatically and applied to the algorithm as a temporal constraint. However the exact morphology of the reference signal in cICA algorithm is relatively unimportant and there is no need for exact similarity between reference and desired signal. E.g. in repetitive transients as our work, using square pulses in the interest region is sufficient. Here we obtain the reference signals by applying proper thresholds on channels T8 and F4. After performing the cICA, RAP-MUSIC algorithm is applied to the epileptic subspace spanned by spatial components of epileptic patterns. We have used the same three layer model (scalp- skull and brain) and the same electrodes as used in simulating data for source localization. Also for constructing a 3D MUSIC map, a score s (which is the estimation of maximum correlation between signal subspace and the potential distribution of source in a defined location), must be obtained. Plotting a function of s, $\log_{10}[1/(1-s^2)]$, can form a 3D MUSIC map in which the maximum point, is the point of a best correspondence between dipolar field distribution and spatial information of epileptic spikes.

IV. Results

A. Source separation

Our simulated data has 15 channels as seen in fig.1 with 1500 time points on which 40 epileptic patterns are superimposed (20 spike and 20 spike-wave patterns). Besides, one of the channels in the underlying EEG data has 50 Hz noise due to the disconnection of electrode. We use it to assign the effect of noise as an independent component.

By applying the cICA algorithm to this data using the reference signals obtained by simple thresholds on F4 and T8 channels (the former is a reference for spike component and the latter is for spike-wave component), two components are extracted as shown in fig. 2. One of them is spike component and the other is the spike- wave component. As seen clearly the algorithm is capable of isolating components due to the applied reference signal. Also the topographic maps of these ICs are depicted in fig. 3. The spatial distributions of topographic maps are too close to the location of simulated epileptic dipolar generators.

The spike-wave component at the output (the above panel of fig. 2) has 19 transients out of 20 but the spike component (the lower panel) has only 16 spike transients in addition to 4 spike- wave transients. Also the spike component has some underlying noise, i.e. the spike component is not pure and is a mixture of multiple underlying components.

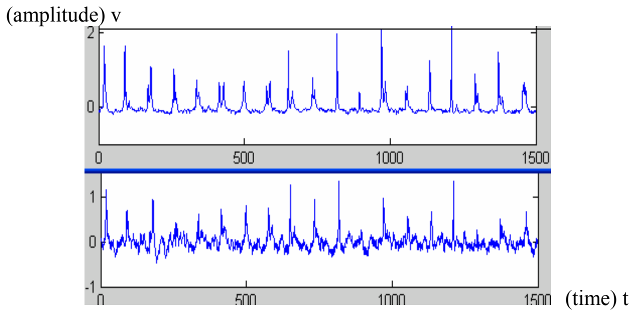

Fig. 2. cICA output after applying the algorithm to the simulated data. Two components, spike-wave (the above panel) and spike (the lower panel) are extracted by using corresponded references.

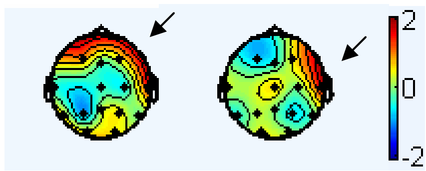

Fig. 3. Topographic maps for recovered ICs. The right figure is for spike-wave component and the left one is for spike component.

It is important to test the algorithm performance (sensitivity) due to false references. For several trials, given a false reference, the algorithm did not converge to the false reference. Also we test the algorithm with a 50 Hz noise as a reference signal in which the noise component is appeared at the output with the conformity to the place of the electrode disconnection.

The accuracy of the recovered ICs compared to the sources are expressed using the peak signal to noise ratio (PSNR) in dB: $\mathrm{PSNR} = 10\log_{10}\sigma^2/\mathrm{MSE}$ where σ^2 denotes the variance of the signal and MSE denotes the mean square error between the original and extracted signal. Also the performance of the source separation is measured by a performance index (PI):

$$\mathbf{PI}_i = \sum_{j=1}^{n} \frac{|p_{ij}|}{\max_k |p_{ik}|} - 1 \qquad (5)$$

Where p_{ij} is the jth element of the vector $\mathbf{p}_i = \mathbf{A}^T\mathbf{w}_i$ [2].

PI is zero when all ICs are perfectly separated.

Here the accuracy and effectiveness of the algorithm is compared with the two conventional ICA algorithms (FastICA and nonlinear PCA[3]). As seen in Table 1 the present algorithm is superior over those mentioned analysis with higher signal to noise ratio and lower PI (which means a better separation).

TABLE I

COMPARISON OF THE RESULTS OF SPIKE COMPONENT RECOVERING FROM A MIXTURE DATA BY USING cICA, FastICA and nonlinear PCA. cICA HAS HIGHER PSNR AND LOWER SEPARASION ERROR

algorithm	PSNR(dB)	PI
PCA	24.45	0.57
Fast ICA	31.08	0.41
CICA	35.12	0.38

[3] Principal Component Analysis

B. Source localization

The single dipoles for the field distributions of the epileptic ICA components are shown in fig. 4 and fig. 5 which were obtained by RAP- MUSIC algorithm. As seen, the estimated dipolar locations of the epileptic ICA components were superimposed on an average MRI scan. The dipolar source estimated for spike component (fig. 4) is close to the simulated spike dipole (EA1 in fig. 1) with the mean 16.5 mm spatial distance and 4° angular difference. These quantities for spike-wave component (EA2 in fig. 2) are 8.3mm and 6.5° respectively.

Finally in order to evaluate the method and determine the sensitivity and specificity of it, the true positive rate vs. false positive rate is computed. The false positives are which further than some threshold – e.g. 100ms- from the nearest true spike. These performance values are plotted as ROC (Recursive Operating Characteristic) curves and both the performance of cICA and dipolar localization algorithms are evaluated. Note that different points of curves are generated by varying the threshold. The two ROC curves corresponding to the detecting and localization stages are shown in fig. 6. The ideal ROC curve would show 100% sensitivity at the lowest possible false positive rate, i.e. the ideal curve rises rapidly towards a true positive fraction approaching unity.

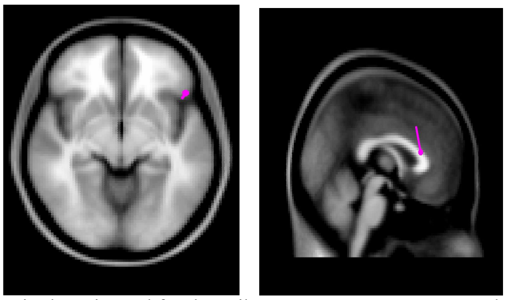

Fig. 4. Dipole estimated for the spike-wave component extracted from a mixture dada by cICA

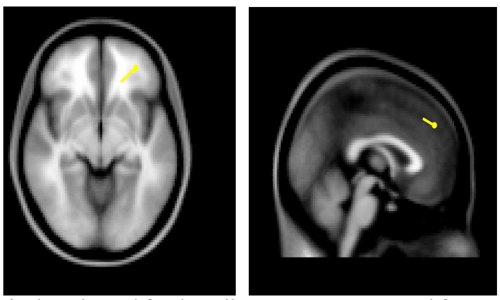

Fig. 5. Dipole estimated for the spike component extracted from a mixture dada by cICA

IV. Conclusion

Conventional ICAs have the disadvantages that extracted components are not ordered and some manual identification of the component(s) of interest is therefore necessary. Also without dimensionality reduction a conventional ICA extracts many components from a large number of measurement channels.

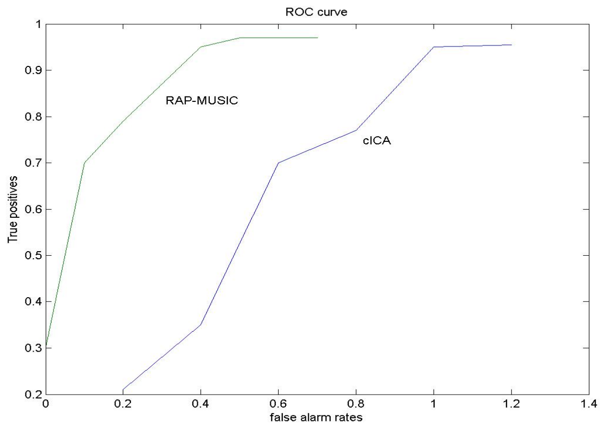

Fig. 6 Recursive Operating Characteristic (ROC) curve showing performance of the two stages (detection and localization).

The cICA method extracts only one single component which is closest by some measure, to a supplied reference. As seen the algorithm did not respond to the false reference, i.e. with multiple false references the algorithm does not extract the false component. Also as illustrated in the last section, this algorithm can extract independent components much purely. Specially, for the spike and wave pattern, which is dissimilar to the sharp brain waves and other artifacts as EOG and EMG, the results are superior. cICA has computational advantages too and it repeatedly converges to the desired component within a few iteration(approximately 9 iterations). Finally it is concluded that the attempt for localizing current dipoles to account for each epileptic component with spatial information achieved by cICA, has improved performance due to the high specificity of algorithm.

References

[1] R. Vigario, "Independent component approach to analysis of EEG and MEG recordings, "*IEEE Trans. Biomed. Eng*., vol. 47, no. 5, May 2000.

[2] W. Lu, and C. Rajapakse, "Approach and application of constraint ICA,"*IEEE Trans. Neural Network,* vol.16, no.1, pp. 203-212 January 2005.

[3] C.James, O. Gibson, "Temporally Constrained ICA: an application to artifact rejection of electromagnetic brain signal analysis," *IEEE Tran. Biomed. Eng*., vol. 50, no. 9, pp. 1106-1114, 2003.

[4] A. Hyvarinen, "fast fixed point Anlgorithm for Independent component analysis," *IEEE Trans. Neural Network*, vol. 10, no.3, pp. 626-634, 1999.

[5] K. Kaboyashi, C. james, and T. Nakahori "Isolation of epilepthiform dishcharge from unaveraged EEG by Independent component analysis," *Clinic.Neurophys*ology, No.100, pp. 626-634, 1999. *Japan,* p. 301, 1982.

[6] J. Mosher, and R. Leahy, "Source localization using recursively applied and projected (RAP) MUSIC," *IEEE. Trans. Sig. Process.,* vol. 47, no.2, pp. 332-340, February, 1999.

[7] N. Acir, I. Oztura, M. Kuntalp, and C. Guzelis, "Automatic detection of epileptiform events in EEG by a three stages procedure based on artificial neural networks," *IEEE Trans. Biomed.* Eng.*,* vol.52, no.1, pp.30-39, January 2005.